

(19) 世界知识产权组织
国际局(43) 国际公布日:
2004年2月19日(19.02.2004)

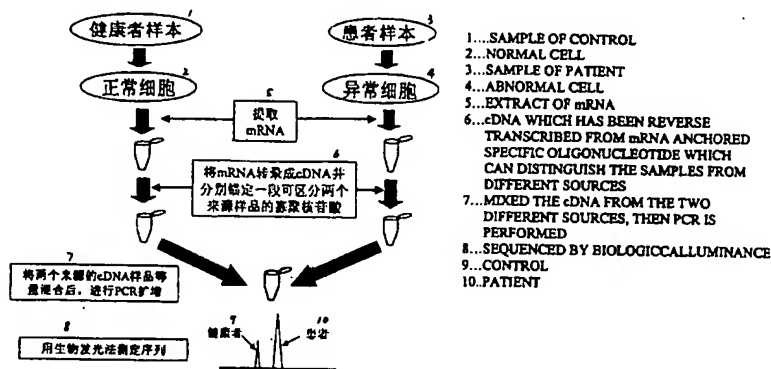
PCT

(10) 国际公布号:
WO 2004/015137 A1

- (51) 国际分类号: C12Q 1/68
- (21) 国际申请号: PCT/CN2003/000648
- (22) 国际申请日: 2003年8月8日(08.08.2003)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权: 02138093.7 2002年8月9日(09.08.2002) CN
- (71)(72) 发明人/申请人: 周国华 (ZHOU, Guohua)
[CN/CN]; 中国江苏省南京市中山东路293号, Jiangsu 210002 (CN).
- (71) 申请人(对除美国以外的所有指定国): 华东医学生物技术研究所(HUADONG RESEARCH INSTITUTE FOR MEDICINE AND BIOTECHNICS) [CN/CN]; 中国江苏省南京市中山东路293号, Jiangsu 210002 (CN).
- (74) 代理人: 北京金信联合知识产权代理有限公司(KINGSOUND & PARTNERS); 中国北京市西城区三里河东路5号中商大厦502室, Beijing 100045 (CN).
- (81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW
- (84) 指定国(地区): ARIPO专利(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- 本国际公布:
— 包括国际检索报告。
- 所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: METHOD FOR COMPARING GENE EXPRESSION LEVEL

(54) 发明名称: 基因表达量比较分析法



(57) Abstract: The present invention provides a method for comparing and determining the gene expression level, which is important in disease-related gene screening, clinical early diagnosis as well as medicine development. Biologicalluminescence process characterised in quantitative analysis was used in the present invention to compare the gene expression level of different individuals or samples and to search disease-related genes. The method of present invention included: (i) After tagged by proper method, the mRNA from different sources were mixed equally, then used as substrate of PCT; (ii) PCR amplification was performed with primers corresponding to the various resources of the genes and primers specific for the genes; (iii) the sequences were analyzed by bioluminescence analysis process, in which, resources of genes were represented by various bases and the expression level of genes were represented by signal intensity.

[见续页]

**(57) 摘要**

本发明涉及一种用于定量和比较基因表达量的分析方法，该方法对疾病相关基因的筛选、临床早期诊断和药物开发具有重大意义。本发明的方法利用生物发光法的定量特性，通过比较同一基因在个体间或样本组间的基因表达量，从而寻找疾病相关基因。具体步骤为：(i)将不同来源的信使核糖核酸(mRNA)用适当的方法标记后等量混合，作为聚合酶链式反应(PCR)的底物；(ii)用与基因各来源相对应的引物和基因特异的引物进行 PCR 扩增；(iii)用生物发光分析法测定序列，以碱基种类代表基因的不同来源，信号强度代表各来源中基因的表达量。

基因表达量比较分析法

技术领域

本发明涉及一种定量测定 DNA 片段混合物中各 DNA 片段含量的方法，具体说是一种用生物发光法同时测定各来源基因表达量的方法，可用于各来源基因表达量的比较分析。

背景技术

近年来随着分子生物学的进步，已测出几十种生物的全基因组顺序，人类基因组计划也即将完成^[1,2]。基因组计划的第一步是完成基因组的构造解析，第二步是进行基因功能解析，即了解基因转录产物 mRNA 的分布状态及 mRNA 的翻译产物蛋白质在细胞及脏器组织中的存在量、分布情况和作用。通过比较基因转录产物 mRNA 的表达量就可以推定未知基因的功能、解明基因间及蛋白质间的相互作用^[3]。因此基因表达量的分析已逐渐成为 DNA 分析的主要课题，在临床医学中，通过比较基因在不同个体（健康和疾病）（脑、肝或心脏）中的表达量，可以寻找和发现疾病基因，进而创造可治疗疾病的特效药，在疾病预防中因为肿瘤、糖尿病和肥胖病等是多基因疾病，用通常的方法很难做到及时诊断，但通过测定与疾病相关的基因在有关组织器官中的表达谱就早达到及可预见患病的危险性，以达预防的目的。此外在分子生物学研究中，通过基因表达量的分析，可以寻找功能基因，为生物品种的改良提供依据。到目前为止可用于基因表达量的方法有：Northern 印迹法^[4]、RT-PCR 法^[5]、测序法^[6,7]、微阵列法^[8,9,10]（microarray）等。

Northern 印迹法是经典的方法，主要用于数个或数十个基因的表达量分析。但该法测定极其繁琐，操作人员必须经过特殊训练，且需使用放射性物质，对人和环境有危害，测定效率不高。

此外该法的灵敏度也很低，不能检出微小的基因表达量。

RT-PCR 法是将 mRNA 反转录成 DNA 后，用各基因的特异性引物与多聚腺嘌呤进行 PCR 扩增反应，然后用电泳分离各 DNA 片段，根据各样品的电泳条带的深浅判断有关基因表达情况和各基因表达量。该法灵敏度虽高，但重复性差，虽然内标可以定量，但定量性不好，因为 PCR 产物量与 DNA 底物量之间的线性关系差，测定结果不能很好地反映实际基因表达情况。

测序法是通过大规模测定 cDNA 的碱基序列而计算基因表达量的方法。主要有整体测序法 (Body Mapping 法) 和基因表达系列分析法 (Serial Analysis of Gene Expression, 简称 SAGE 法)。该法结果准确可靠，前一种方法是用测序仪分别测定代表各基因的 cDNA 序列在样品中的出现频率，工作量很大，且需要许多昂贵的仪器；后一种方法，是经改进的测序法，即先用限制性内切酶把固定于微球 (bead) 表面的 cDNA 切成带有粘性切口的片段，然后，将与微球相连的 DNA 分成二等分，分别加入带有粘性切口的连接 DNA A 和 B (记为 Linker-A 和 Linker-B)，并使之连接，然后再用 $\text{H} \text{S}$ 内切酶 (标记酶) 切成带有 9~12 个标记碱基的片段，标记片段是用于鉴别基因的短核苷酸序列，因为 $\text{H} \text{S}$ 型内切酶能从特异性酶切位点开始切割成一定长度的标记碱基片段。将含有 Linker-A 和 B 的两部分混合，并用连接酶连接。加入引物 A (Linker A 中的部分) 和引物 B (Linker-B 中的部分) 进行 PCR 扩增，PCR 产物中含有两个基因特异性标记。再用上述内切酶将 PCR 产物切割成含有 4 个碱基粘性切口的 cDNA 片段，连接后进行克隆，每个克隆中含有 10~50 个基因标记片段，每两个标记之间有一个含有四个碱基的短片段 (内切酶的特异性序列) 相隔。最后测定各克隆的序列，根据各标记在序列中的出现的频率计算各基因的表达量。该法虽然不需分别测定各 cDNA 的序列，但工作量仍然很大，且操作较为繁琐，需要使用价格昂贵

的测序仪，普通实验室无法做到，非常规测定方法。

微阵列法是在固相表面上连接寡聚核苷酸片段（20~30个碱基）或 cDNA，将生物样品中的 mRNA 提取转录成 cDNA 后与之杂交，同一块芯片可与两种不同来源的样品杂交，每一样品标记有不同荧光基团，通过芯片杂交点颜色的差异比较两种不同来源样品中 mRNA 的相对含量。该法的缺点是灵敏度低，定量性差，需要专门软件处理数据，且仪器价格昂贵。

焦测序法^[11]是一种采用生物发光的方法测定 DNA 序列，仅能准确测定 10~30 个碱基序列，主要用于 SNP 分析中^[12]。该法定量性能好，灵敏度高，操作简便，不需电泳，不需标记反应，测定时不需激光源，也无需使用特殊或昂贵的试剂，所需仪器简单。焦测序法是一种基于用发光分析法测定焦磷酸盐（PPi）的测序技术，测定原理如下：

在核酸聚合酶的作用下，如果加入互补的 dNTP，与引物退火的单链 DNA 就发生延伸反应，释放出等摩尔量的 PPi；PPi 在三磷酸腺苷硫酸化酶（ATP sulfurylase）的催化作用下转化为三磷酸腺苷（ATP）；萤光素酶（luciferase）催化作用下，ATP 与虫萤光素（luciferin）反应发光，可用光电倍增管（PMT）或电荷耦合装置（CCD）检测。由于光信号强度与 PPi 的量成正比，所以根据加入的 dNTP 种类和信号强度就可确定靶 DNA 的序列。

但该法不能直接用来比较各来源的基因表达量。

发明内容

本发明的目的就是为了解决上述问题，提出一种灵敏度高、定量性好、价格低廉、操作方便的基因表达量比较分析法。

本发明的技术解决方案：

一种基因表达量比较分析法，其特征在于它具有如下特征：

（a）将不同来源的信使核糖核酸（mRNA）用适当的方法标记后，等量混合，作为聚合酶链反应（PCR）的底物；

(b) 用与基因各来源相对应的引物和基因特异性引物进行 PCR 扩增;

(c) 用生物发光分析法测定序列, 以碱基种类代表基因的不同来源, 信号强度代表各来源中基因表达量。

所述“不同来源的 mRNA”是指同一基因在同一生物种的不同个体或同一基因在同一个个体的不同器官, 如健康人和病人某一器官中的 mRNA; 也可指同种个体中的同一基因在不同的化学或物理刺激下产生的表达。

所述“与各来源相对应的特异性引物”是指一组序列不同, 但由相同种类和数目的碱基所组成的引物, 每一种引物代表一种基因的来源。

所述“用适当的方法标记”是指用适当长度的 DNA 片段来区别各来源基因的方法, 第一种方法: 将不同来源的 mRNA 用反转录 PCR (RT-PCR) 法制成双链互补 DNA (cDNA) 后, 用限制性内切酶将之切成适当长度的片段, 用选择性适配器 (adapter) 与之相连接, 不同来源的 mRNA 连有不同的 adapter; 第二种方法: 将带有多聚胸腺嘧啶引物的微球分别与不同来源的 mRNA 杂交, 并制成第一链 cDNA, 用 5'端含有与各来源相对应序列的锚定引物进行互补链 (第二链 cDNA) 的合成, 锚定引物的 5'端用于区别基因的不同来源; 第三种方法: 直接用 5'端含有与各来源相对应序列的锚定引物与 mRNA 杂交, 并制成第一链 cDNA, 锚定引物 5'端的构造与第二种方法相同。

本发明解决了现有用于比较个体之间的 mRNA 表达量的方法均存在不同的缺点及不易推广使用的问题, 主要借助于焦测序法的原理, 用生物发光的方法定量测定和比较不同来源样品中的 mRNA 表达量, 达到寻找与疾病相关基因的目的, 并用于临床诊断。本发明具有灵敏度高、定量性好、价格低廉、操作方便的优点。

附图说明

图 1 是本发明的测定原理图。

图 2 是酶解法测定两个体中的基因表达量示意图。

图 3 是 DNA 适配器的构成示意图。

图 4 是反应模块示意图。

图 5 是生物发光法测定两个不同来源基因表达量的图谱。

图 6 是使用 96-孔反应板同时测定多基因表达量仪器装置示意图。

图 7 是两组混合样品中平均基因表达量的比较示意图。

图 8 是用不含有三磷酸腺苷双磷酸酶的 PPi 测定溶液测定实施例 1 中的样品结果示意图。

具体实施方式

下面结合附图对本发明作进一步说明。

本发明方法由三步组成：(1) 各来源 cDNA 的转录和标识；(2) 等量混合各来源的 cDNA，并作为 PCR 的扩增模板；(3) 用生物发光法测定序列。测定原理如图 1 所示。

本发明第一个关键点是如何在 PCR 扩增前用适当的方法标识各来源的 cDNA，而又不影响 PCR 等比例扩增。有几种技术可实现这一方案：(1) 将 mRNA 反转录成双链 cDNA 后，用限制性内切酶消化成若干片段，再用与可识别各 cDNA 来源的 DNA 适配器连接，各 DNA 适配器由相同碱基组成，但序列不同，连接后将各来源的 cDNA 样品等量混合后，进行 PCR 扩增；(2) 用固定在微球上的多聚胸腺嘧啶与 mRNA 杂交，然后再进行 cDNA 合成，cDNA 第一条链合成完后，加入 5' 端含有锚定序列的基因特异性引物进行互补链的合成，引物 5' 端的附加锚定碱基序列用于区别各 cDNA 的来源。cDNA 合成完后，将该 DNA 链与微球分离，并作为 PCR 扩增的模板。最后用基因特异性引物和区别各来源的锚定序列部分作为引物进行 PCR 扩增；(3) 用 5' 端带锚定碱基序列的基因特异性引物与 mRNA 杂交并进行 cDNA 第一条链的合成，其中引物中的锚

定碱基序列部分与基因各来源相对应，各来源单链 cDNA 等量混合后，进行 PCR 扩增，PCR 扩增的引物为锚定碱基序列中用于区别基因各来源的部分。

本发明第二个关键点是如何将焦测法所测结果中的序列变为各来源的信息，信号强度变为各来源 mRNA 的表达量。本发明是通过如下方法来实现这一目的的，即在 PCR 扩增前，在 cDNA (PCR 扩增模板) 中引入序列不同的数个碱基，但所引入的碱基总数和种类相同，仅序列不同而已，这样混合后进行 PCR 反应可实现等比例扩增，最后加入与各 cDNA 来源相对应的引物混合物进行测序反应，不同序列即代表不同的来源，峰高代表表达量。以下用例子来说明。

实施例 1，两个体中基因表达量的比较：

本例叙述了用限制性内切酶法切割 cDNA 片段后与 DNA 适配器连接，然后等量混合来源-A 和来源-B 的 cDNA 进行 PCR 扩增的方法，其测定过程如图 2 所示。以人 P53 基因为例进行了说明，其中 mRNA 的提取采用 Gibico TRIzol-LS-Reagent™ 试剂盒的标准方法。

1. cDNA 样品的制备 (各来源的 cDNA 样品分别制备)

取 1 μg 来源-A 或来源-B 的 mRNA 样品，加入 0.5 μg Oligo(dT)₁₂₋₁₈，5.5 μl H₂O，并混合均匀，采用 cDNA 第一链合成试剂盒 (Gibico Super Script™ Preamplification System for First Strand cDNA Synthesis)，在 0℃ 下保温 10min 后，加入预先制备好的混合液 (4 μl 的 5 倍缓冲液 (I)，2 μl 的 0.1 M 二硫苏糖醇，1 μl 的 10 mM dNTPs，2 μl H₂O 和 1 μl 的 200 U/ μl 反转录酶)，于 42℃ 下保温 50min，之后在 70℃ 下保温 15min，反应后保存于 10℃。

在上述反应混合液中进一步加入 10 μl 10 倍连接酶缓冲液，70 μl H₂O，1 μl 1 mM dNTPs 混合液，50U 聚合酶 I，10U DNA

连接酶，搅拌后加入 2U 核糖核酸酶 RNase H，混匀。在 10℃ 下保温 2hr，然后于 70℃ 下保温 15min，该混合物即为双链 cDNA 样品，于 -20℃ 下冷冻保存。

取上述制备好的 cDNA 溶液 150 μ l，加入 10 倍限制性内切酶 *Mbo* I 缓冲液 (200 mM Tris-HCl (pH8.5)，100 mM MgCl₂，10 mM DDT 和 1000 mM KCl) 20 μ l，灭菌蒸馏水 30 μ l，总量为 150 μ l。加入 60 U 的限制性内切酶 *Mbo* I，在 37℃ 的条件下保温反应一夜，反应完后，于 75℃ 的条件下，保温 15min，使 *Mbo* I 酶失活。

2. DNA 适配器的制备

本例中，使用 DNA 适配器来区别不同来源的基因表达量。首先按上述方法用 *Mbo*I 限制性内切酶将双链 cDNA 切割成 5'端突出四个碱基“ctag”的楔型结构。然后将 DNA 适配器与之相连接。在图 3 中，“Y”型 DNA 适配器由 2 条部分互补的链组成，链“a”的 5'端臂用以区别各基因的来源，每个 5'端臂含有相同数目和种类的碱基，但排列序引不同；链“b”的 3'端翘起，即与链“a”不互补，目的是在 PCR 扩增反应时，链“b”不会发生延伸反应，其 5'端有四个突出的碱基，为酶切位点，序列为“gatc”，3'端磷酸化。

分别用 DNA 适配器-A 和-B 来识别基因来源 A 和来源 B，它们的区别主要是链“a”的 5'端部分，其余序列相同。

设 P-1 为适配器-A 的“a”链，其序列为：

P-1: 5'-ccc cac ttct tggt ctc tcat cagg cgc atc act cg-3'

设 P-2 为适配器-B 的“a”链，其序列为：

P-2: 5'-cac ctc tcat ttct ccc tggt gacg ogc atc act cg-3'

适配器-A 和-B 链的“b”链相同，用 P-3 表示，其序列为：

P-3: 5'-gatccg agt gat gcg ct aag-3'

在来源 A 的酶解 cDNA 溶液中，加入 10 pmol 的 P-1 和 P-3，在来源 B 的酶解 cDNA 溶液中加入 10 pmol 的 P-2 和 P-3。并含有

pH7.6 的 Tris-HCl, 6.5 mM MgCl₂, 0.5 mM ATP, 0.5 mM DDT 和 2.5% 聚乙烯甘油-800, 在 70℃ 的条件下, 反应 10 min, 然后慢慢冷却到 16℃。加入 T₄ DNA 连接酶, 在 16℃ 下反应 2h。连接产物为下一步 PCR 反应的底物。

3. PCR 扩增 cDNA 限制性酶切片段及单链的制备

将来源 A 和 B 的 cDNA 限制性内切酶片段与适配器的连接产物等量混合, 作为 PCR 反应的底物。以适配器 5' 端部分为 PCR 反应引物, 即由 P-1 和 P-2 中的前 21 个碱基组成的引物 (以 5' 端计算) 混合物, 并分别记为 MP-1 和 MP-2。另外选取 P53 基因中的一段序列作为特异性引物, 记为 GSP, 并且在 5' 端用生物素 (biotin) 标记。取等量的 MP-1、MP-2 和 GSP 混合后作为 PCR 引物。在 10 μl PCR 反应混合物中, 底物为 1 μl, 引物均为 1 pmol, 缓冲液为 pH8.0 的 20 mM Tris-HCl, 50 mM KCl, 0.2 mM dNTPs 和 1.25 U DNA 聚合酶。PCR 反应条件为: 94℃ 30s, 58℃ 1min, 72℃ 30s, 反应 30 个循环。最终得到的产物是生物素标记的双链 DNA。将其与抗生物素蛋白链菌素 (streptavidin) 涂层微球 (Dynabeads M280) 于室温下反应 30 min, 反应介质为 5 mM Tris-HCl, pH7.5, 0.5 mM EDTA, 1.0 M NaCl, 除去上清后加入 0.1 M NaOH 溶液, 反应 5 min, 冲洗干净后, 在微球冲洗缓冲液 (5 mM Tris-HCl (pH7.5), 0.5 mM EDTA 和 1.0 M NaCl) 中储存备用。该产物为来源 A 和 B 的 DNA 单链混合物。

4. 生物发光法测定各来源基因表达量

取上述单链 DNA 样品 (步骤 3 中的微球) 制备成含有 25 mM Mg²⁺ 和 5 mM Tris (pH 7.7) 的溶液, 并分别加入 5 pmol 的 MP-1 和 MP-2, 94℃ 加热 2min 后自然冷却至室温。取 1~5 μl 加入到 50~100 μl 的 PPI 测定标准混合液, 依次加入 dGTP 和 dCTP, 或 ddGTP 和 ddCTP, 或它们的类似物, 进行测序反应。加入 dGTP 时所得信号强度代表来源 A 的基因表达量, 加入 dCTP 时所得信号强度代

表来源 B 的基因表达量。

PPi 测定标准混合液组成为：0.1 M Tris-HAc (pH 7.7), 2 mM EDTA, 10 mM Mg (Ac)₂, 0.1% 白蛋白 (BSA), 1mM 二硫苏糖醇 (DTT), 3 μ M 5' -磷酸化硫酸腺苷 (APS), 0.4 mg/ml 聚乙烯吡咯烷酮 (PVP), 0.4 mM 荧光素, 200 mU/ml 三磷酸腺苷硫酸化酶, 2U/ml 三磷酸腺苷双磷酸酶 (apyrase), 1U 无外切酶活性的 DNA 聚合酶 Klenow, 及适量荧光素酶。

5. 测定仪器装置

按图设计可供一个样品的测定装置, 本发明的仪器关键部分为反应模块, 如图 4 所示。在图中正中间的反应池底部通过毛细管分别连接两个 dNTP 贮液槽, 当 dNTP 贮液槽上方施以一下压力时, dNTP 或 ddNTP 就通过底部的管道 (毛细管) 流向反应池。反应产生的荧光透过底部材料进入检测器。检测器可以用光电倍增管、光电二极管或电荷耦合照相机 (CCD)。

6. 测定结果

在图 4 的反应的模块中, 分别加入 dGTP 和 dCTP 于两个 dNTP 贮液槽中, 将样品和 PPi 测定标准混合液加入中间反应池中, 分别在 dGTP 和 dCTP 贮液池的上方用注射器加压, 反应产生的信号如图 5。根据信号强度, 可以计算出两个来源基因的相对表达量。

实施例 2, 用 96-孔反应板同时测定 96 个不同基因的表达量:

本实施例主要是采用普通 96-孔反应板, 同时测定疾病组与健康组间的 96 个不同基因的表达量。

1. 测定样品的制备

按实施例 1 的方法, 提取来源 A 和来源 B 的 mRNA, 并制成双链 cDNA, 然后用限制性酶 Mob I 切割, 再加入两种分别代表来源 A 和 B 的 DNA 适配器, 用连接酶连接好后, 分别于 96-孔反应板中的每一个孔中, 加入 1~5 μ l 连接好的上述溶液作为 PCR 扩增的底物, 最后加入 MP-1、MP-2 和相应的各基因 GSP 引物进行扩

增反应, GSP 引物的 5' 端用生物素修饰, 按实施例 1 相同的方法制备单链 DNA, 然后在 96 孔中的每一个小孔中加入相同量的 MP-1 和 MP-2 混合物作为测序反应的引物, 实验方法同实施例 1。

2. 测定装置

本实施例的关键点在于购建可以供 96 个样品同时测定的仪器装置。如图 6 所示, 采用压差进样法, 根据 96-孔反应板的结构用毛细管制作两个液体分送器, 每个分送器有 96 根毛细管作为每个反应孔的 dNTP 加注头。

毛细管的一端与 dNTP 或 ddNTP 贮液池相连, 贮液池在 96-孔反应板的上方, 测定时, 通过升降装置将液体分送器的 dNTP 加注头伸入反应液中, 在压差的作用下, dNTP 或 ddNTP 流入反应池中进行掺入反应, 反应释放的 PPi 在三磷酸腺苷硫酸化酶(ATP sulfurylase)的催化作用下转化为三磷酸腺苷(ATP), ATP 与荧光素及荧光素酶作用发出荧光。用电荷耦合装置(CCD)或光电信增管(PMT)或光电二极管检测 96 个反应孔中发出的光信号。当然本装置还适用于同一基因的多个样品的同时测定。

实施例 3, 六个不同来源基因表达量的比较:

通常采用微陈列芯片仅能测定两种不同来源基因的表达量, 如果增加一种来源, 则需要增加一种不同的荧光标记物, 从而使测定成本增加, 且通常同时不会使用超过四种激发波长不同的荧光标记物。由于本发明采用测序法测定不同来源的基因表达量, 测定的碱基序列代表不同的基因来源, 所以增加样品来源, 不会增加任何测定成本。

由于 dATP 是 ATP 的类似物, 背景高, 对测定产生严重干扰, 虽然可用 dATP 的类似物 dATP α s 代替, 但测定成本增加。本发明避免使用 dATP。当用本法比较两个不同来源基因表达量时, 分别在 P-1 和 P-2 中采用 "cag" 和 "gac" 来控制。因为最多可以加

入“g”、“c”和“t”三种 dNTP 来测序，所以通过加入单个 dNTP 的办法最多只能测定三种不同来源的基因表达量。由于本发明采用的 PPi 测定法具有很好的定量特性，故在 DNA 适配器的中间部分（PCR 扩增引物后）设计如下六种不同的序列用以控制基因的不同来源，每一个序列的第 4 个碱基为“T”，用于控制碱基的进一步延伸。

- | | | |
|-----------|-----------|-----------|
| (1) cgat; | (2) gcat; | (3) agct; |
| (4) gact; | (5) cagt; | (6) acgt; |

除 dATP 外的三种 dNTP 加入顺序为：dTTP、dGTP、dCTP，dTTP 和 dGTP，共加入 7 次，根据测定图谱中 6 个峰的高低计算出各峰所代表的基因表达量。计算方法为解联立方程，由计算机软件自动完成，最终给出各基因来源的相对基因表达量。

测定仪器装置同（实施例 1），但需增加一个 dNTP 贮液池，如测定多个样品可采用（实施例 2）中的仪器装置，但必须增加一个 96 孔加样头，因为有三种 dNTP 需加入到反应池中。

实施例 4，混合样品中平均基因表达量的比较：

本例的目的是测定两组之间，如健康组和疾病组，基因平均表达量的差异。设两组各含有 100 个样本，常规法如基因芯片法，需测定 100 次，然后给出每个样品的测定结果，最后用计算机软件对测定数据进行分析，得出统计结果。本发明将含有 100 个人份的健康组和疾病组的样品分别等量混合成两大组，即健康组和疾病组，然后按照测定两个不同来源样品的方法，测定两组之间的相对基因表达量，该结果为混合样品中的基因平均表达量。然后将两组的基因表达量与疾病的发生、发展相关联，找出与疾病相关联的基因。与基因芯片测定法相比，寻找疾病相关基因的效率提高了两个数量级，芯片法要测定 100 个样本，而本法仅需一次，并且测定结果更为准确。测定方法如图 7 所示。

测定混合样品中平均基因表达量的关键是：（1）用 PCR 法将

两种来源样品中的微量 cDNA 进行等比例扩增；(2) 信号检出应具有很好的定量特性。本发明正好能满足这两点要求，所以可用于比较不同来源混合样品中的平均基因表达量，其测定速度和结果可靠性将超过芯片法。

样品制备采用微球法，即用固定在微球上的多聚胸腺嘧啶与 mRNA 杂交，然后再进行 cDNA 合成，cDNA 第一条链合成完后，加入 5' 端含有锚定序列的基因特异性引物进行互补链的合成，引物 5' 端的附加锚定碱基序列用于区别各 cDNA 的来源，在本例中，用于区别两种基因不同来源的附加锚定序列分别为实施例 1 中 P-1 和 P-2 的前 23 个碱基（从 5' 端计算）。cDNA 合成完后，将该 DNA 链与微球分离，并作为 PCR 扩增的模板。最后用生物素修饰的基因特异性引物 (GSP) 和区别各来源的锚定序列部分（同实施例 1 中的 MP-1 和 MP-2）作为引物进行 PCR 扩增，余下部分同实施例 1。

实施例 5，PPi 标准测定溶中不加入三磷酸腺苷双磷酸酶 (apyrase) 时的基因表达量测定法：

在焦测序法中，因为需要得到连续的测定序列，所以必须加入一种可同时降解 dNTP 和 ATP 的酶“三磷酸腺苷双磷酸酶”，这也是焦测序法的关键技术。本发明中测定两个或三个不同来源的基因表达量时，加入的 dNTP 为三种不同的 dNTP，相互不会产生干扰，即当加后一种 dNTP 时，前一种已加入的 dNTP 不会发生进一步的延伸反应；另外反应溶液中产生的 ATP 也没有必要降解，因为用虫萤光素法测定 ATP 的线性范围很宽，当加入另外一种 dNTP 或两种 dNTP 时，由 ATP 产生的信号强度仍可控制在线性范围内。

本实施例用不含有三磷酸腺苷双磷酸酶的 PPi 测定溶液测定实施例 1 中的样品，得到图 8 的结果，说明不加三磷酸腺苷双磷酸酶是可行的。

文中标引:

- 1 Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, et al.: The sequence of the human genome. *Science* 2001;291(5507):1304-51.
- 2 McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, Sekhon M, Wylie K, et al.: A physical map of the human genome. *Nature* 2001;409(6822):934-41.
- 3 Matsubara, K., and K. Okubo. 1993. cDNA analyses in the human genome project. *Gene* 135: 265-74.
- 4 Kawasaki, E. S., S. S. Clark, M. Y. Coyne, S. D. Smith, R. Champlin, O. N. Witte, and F. P. McCormick. 1988. Diagnosis of chronic myeloid and acute lymphocytic leukemias by detection of leukemia-specific mRNA sequences amplified in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 85: 5698-702.
- 5 Karet, F.E., D. S. Charnock-Jones, M. L. Harrison-Woolrych, G. O'Reilly, A. P. Davenport, and S. K. Smith. 1994. Quantification of mRNA in human tissue using fluorescent nested reverse-transcriptase polymerase chain reaction. *Anal. Biochem.* 220: 384-90.
- 6 Velculescu, V. E., L. Zhang, B. Vogelstein, and K. W. Kinzler. 1995. Serial analysis of gene expression. *Science* 270: 484-7.
- 7 Powell, J. 2000. SAGE. The serial analysis of gene expression. *Methods in Molecular Biology* 99: 297-319.
- 8 Schena, M., D. Shalon, R. W. Davis, and P. O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270: 467-70.
- 9 Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush. 2000. A concise guide to

cDNA microarray analysis. *Biotechniques* 29: 548-50, 552-4, 556 passim.

10 Ferguson, J. A., T. C. Boles, C. P. Adams, and D. R. Walt. 1996. A fiber-optic DNA biosensor microarray for the analysis of gene expression. *Nature Biotechnology* 14: 1681-4.

11 Ronaghi, M., M. Uhlen, and P. Nyren. 1998. A sequencing method based on real-time pyrophosphate. *science* 281: 363, 365.

12 Ahmadian, A., B. Gharizadeh, A. C. Gustafsson, F. Sterky, P. Nyren, M. Uhlen, and J. Lundeberg. 2000. Single-nucleotide polymorphism analysis by pyrosequencing. *Anal Biochem* 280: 103-10.

权 利 要 求

1、一种基因表达量比较分析法，其特征在于它具有如下特征：

(a) 将不同来源的信使核糖核酸 (mRNA) 用适当的方法标记后，等量混合，作为聚合酶链反应 (PCR) 的底物；

(b) 用与基因各来源相对应的引物和基因特异性引物进行 PCR 扩增；

(c) 用生物发光分析法测定序列，以碱基种类代表基因的不同来源，信号强度代表各来源中基因表达量。

2、根据权利要求项 1 所述的基因表达量比较分析法，其特征在于所述“不同来源的 mRNA”是指同一基因在同一生物种的不同个体或同一基因在同一个个体的不同器官，如健康人和病人某一器官中的 mRNA；也可指同种个体中的同一基因在不同的化学或物理刺激下产生的表达。

3、根据权利要求项 1 所述的基因表达量比较分析法，其特征在于所述“与各来源相对应的特异性引物”是指一组序列不同，但由相同种类和数目的碱基所组成的引物，每一种引物代表一种基因的来源。

4、根据权利要求项 1 所述的基因表达量比较分析法，其特征在于所述“用适当的方法标记”是指用适当长度的 DNA 片段来区别各来源基因的方法，第一种方法：将不同来源的 mRNA 用反转录 PCR (RT-PCR) 法制成双链互补 DNA (cDNA) 后，用限制性内切酶将之切成适当长度的片段，用选择性适配器 (adapter) 与之相连接，不同来源的 mRNA 连有不同的 adapter；第二种方法：将带有多聚胸腺嘧啶引物的微球分别与不同来源的 mRNA 杂交，并制成第一链 cDNA，用 5'端含有与各来源相对应序列的锚定引物进行互补链 (第二链 cDNA) 的合成，锚定引物的 5'端用于区别基因的不同来源；第三种方法：直接用 5'端含有与各来源相对应序

列的锚定引物与 mRNA 杂交，并制成第一链 cDNA，锚定引物 5'端的构造与第二种方法相同。

5、根据权利要求项 4 所述的基因表达量比较分析法，其特征在于所述“选择性适配器”是指含有与权项 4 中限制性内切酶切口互补的序列，并在连接酶作用下能与切口完全连接的楔形双链 DNA，其中一条链的 5'端含有权项 3 中与各来源相对应的特异性序列，另一条链的 3'端含有与上述链不互补的碱基，或 3'端经过适当修饰不能在 DNA 聚合酶的作用下发生延伸反应。两条链组成“Y”型的结构，一端不互补，成叉型分开，另一端成权项 4 第一种方法中限制性酶切口的形状。

6、根据权利要求项 4 所述的基因表达量比较分析法，其特征在于所述“选择性适配器”和“锚定引物”中用于区别基因不同来源的部分是指一段碱基种类和数目相同而排列顺序不同的序列，它们具有相同的熔解温度（ T_m 值）。

7、根据权利要求项 1 所述的基因表达量比较分析法，其特征在于所述“生物发光分析法”是指定量测定延伸反应产生的焦磷酸盐（PPi）的方法。

8、根据权利要求项 7 所述的基因表达量比较分析法，其特征在于所述“延伸反应”是指在权项 1 中的 PCR 扩增产物的单链产物中，加入特定的引物或引物混合物进行退火，然后按一定次序分别逐个加入三磷酸脱氧核苷酸（dNTP），或三磷酸双脱氧核苷酸（ddNTP），或它们的类似物，在 DNA 聚合酶的作用下，当所加 dNTP 或 ddNTP 或它们的类似物与模板互补时，所发生的聚合反应。

9、根据权利要求项 8 所述的基因表达量比较分析法，其特征在于所述“单链产物”是指用物理或化学的方法将权项 1 中的 PCR 产物制成的单链产物，其中“物理学方法”指用生物素

(biotin)标记 PCR 中的一种引物,然后用固相法制备单链,“化学方法”是指用酶切法制备单链。

10、根据权利要求项 7 所述的基因表达量比较分析法,其特征在于所述“延伸反应”指权项 1 中的 PCR 扩增产物,经酶处理去除 PCR 反应中的 PPi,过剩的 dNTPs 和过剩的引物后,加入单链结合蛋白(如 SSB),其余按项 8 进行反应。

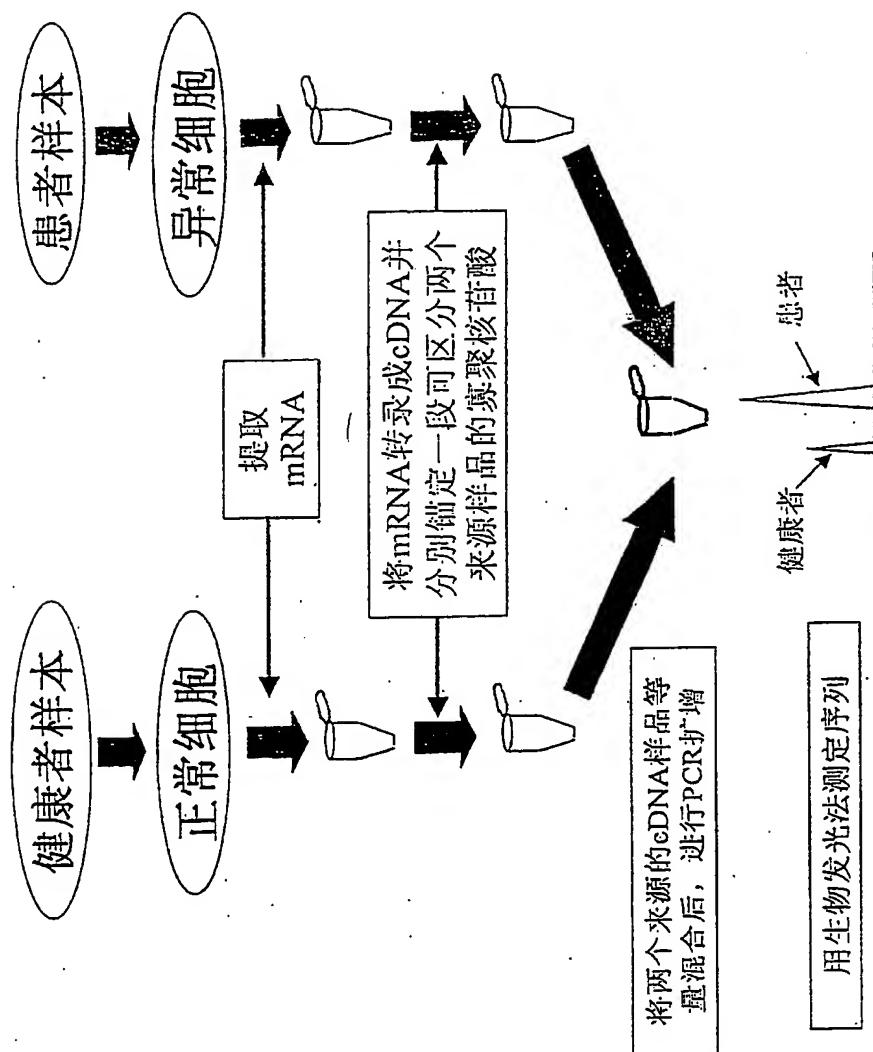


图1

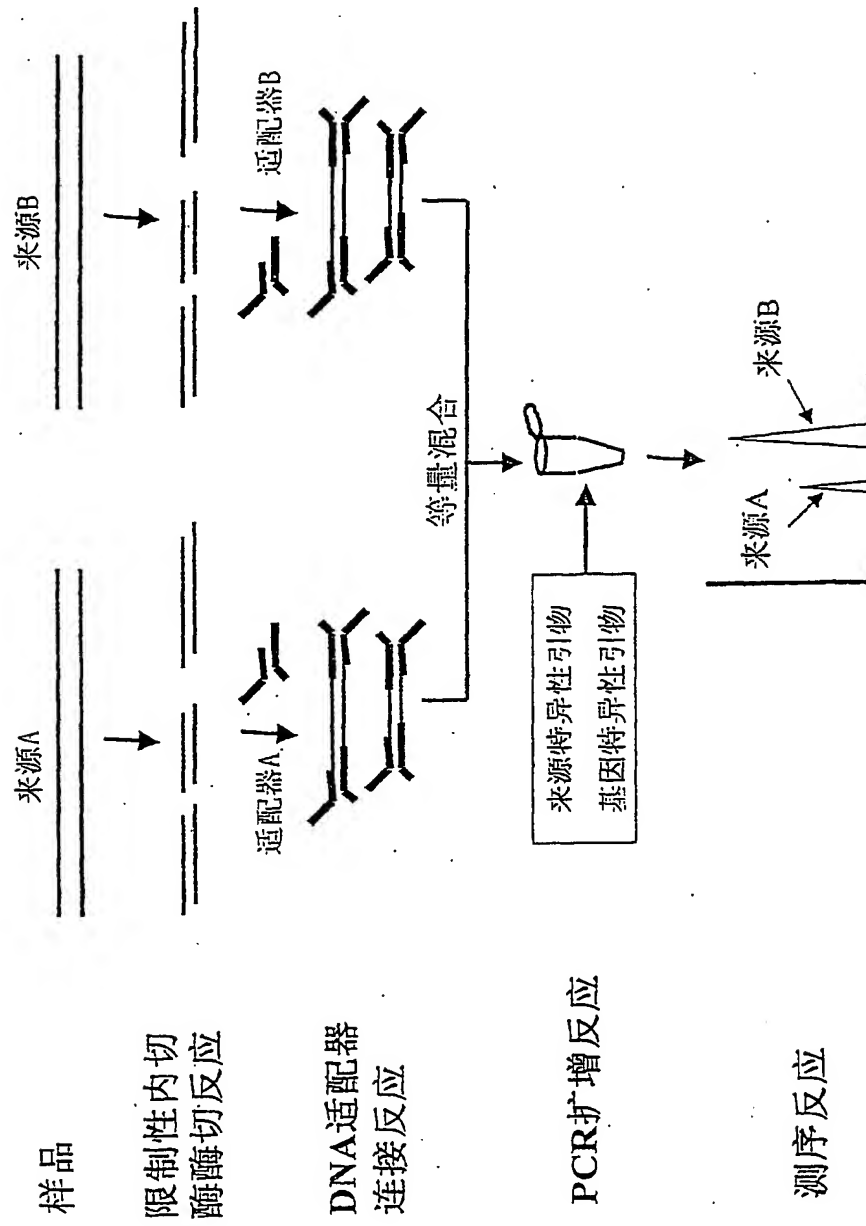


图2

3/8

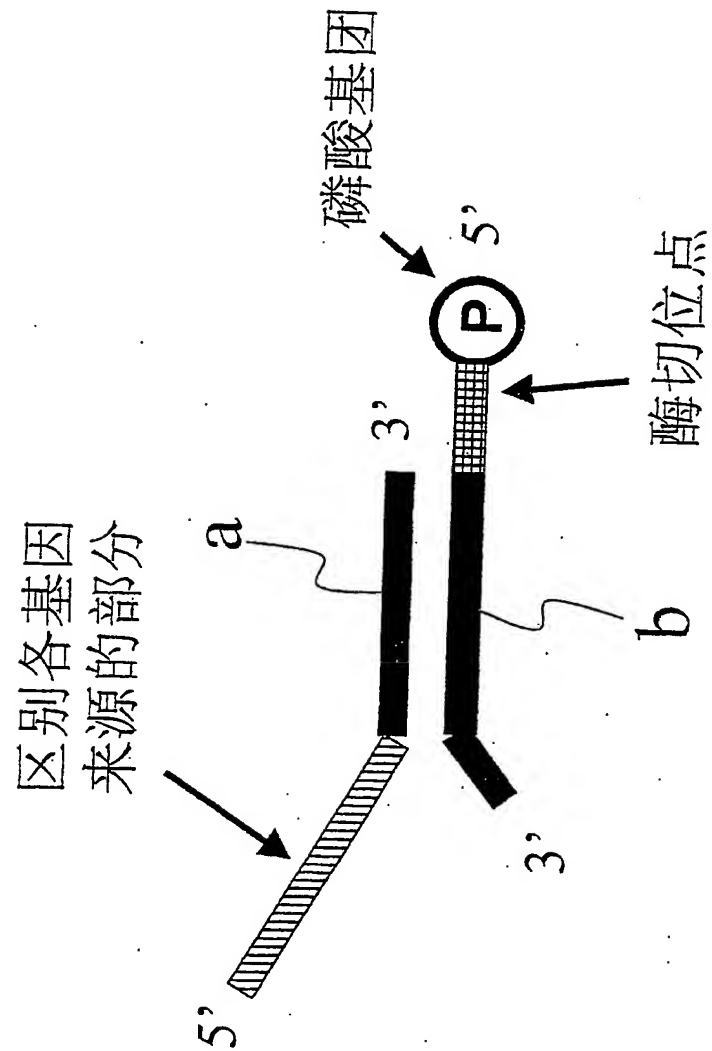


图3

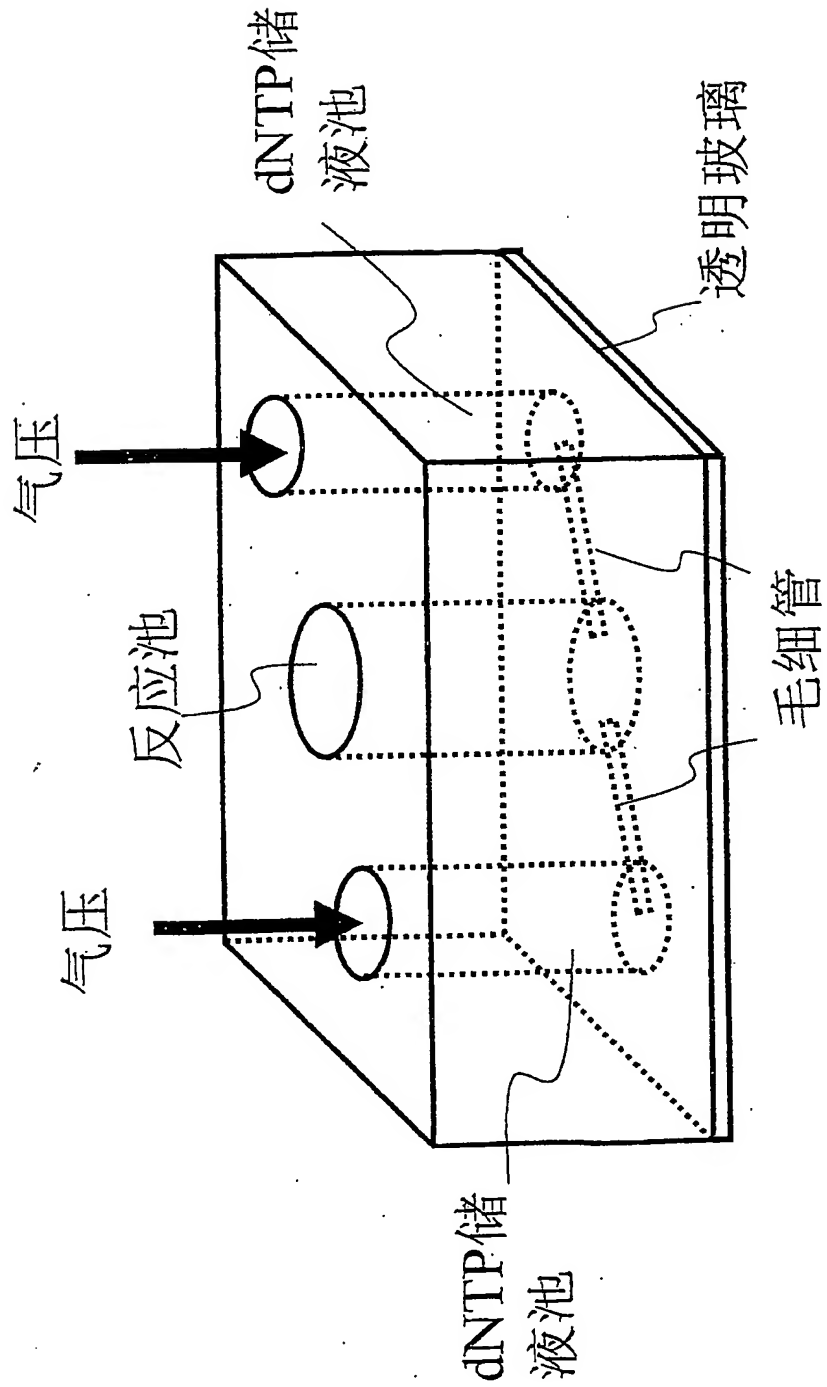


图4

5/8

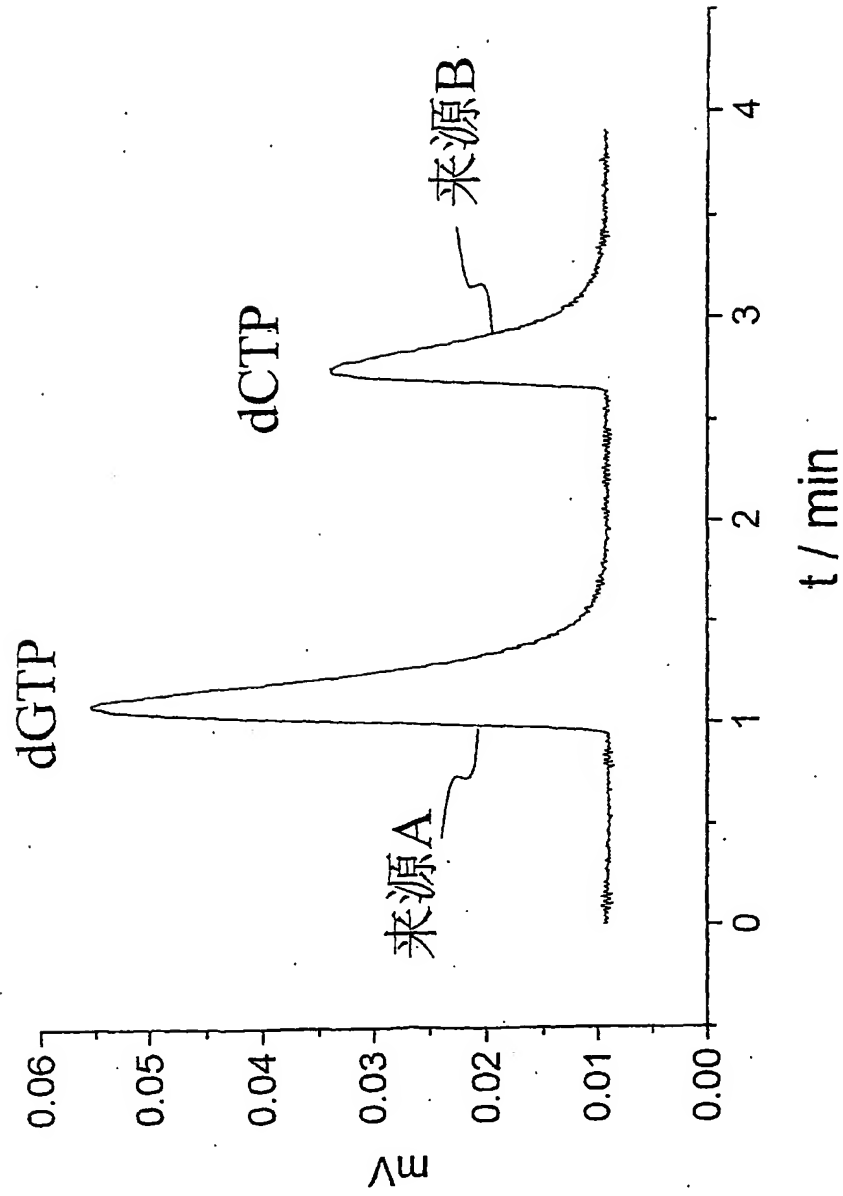
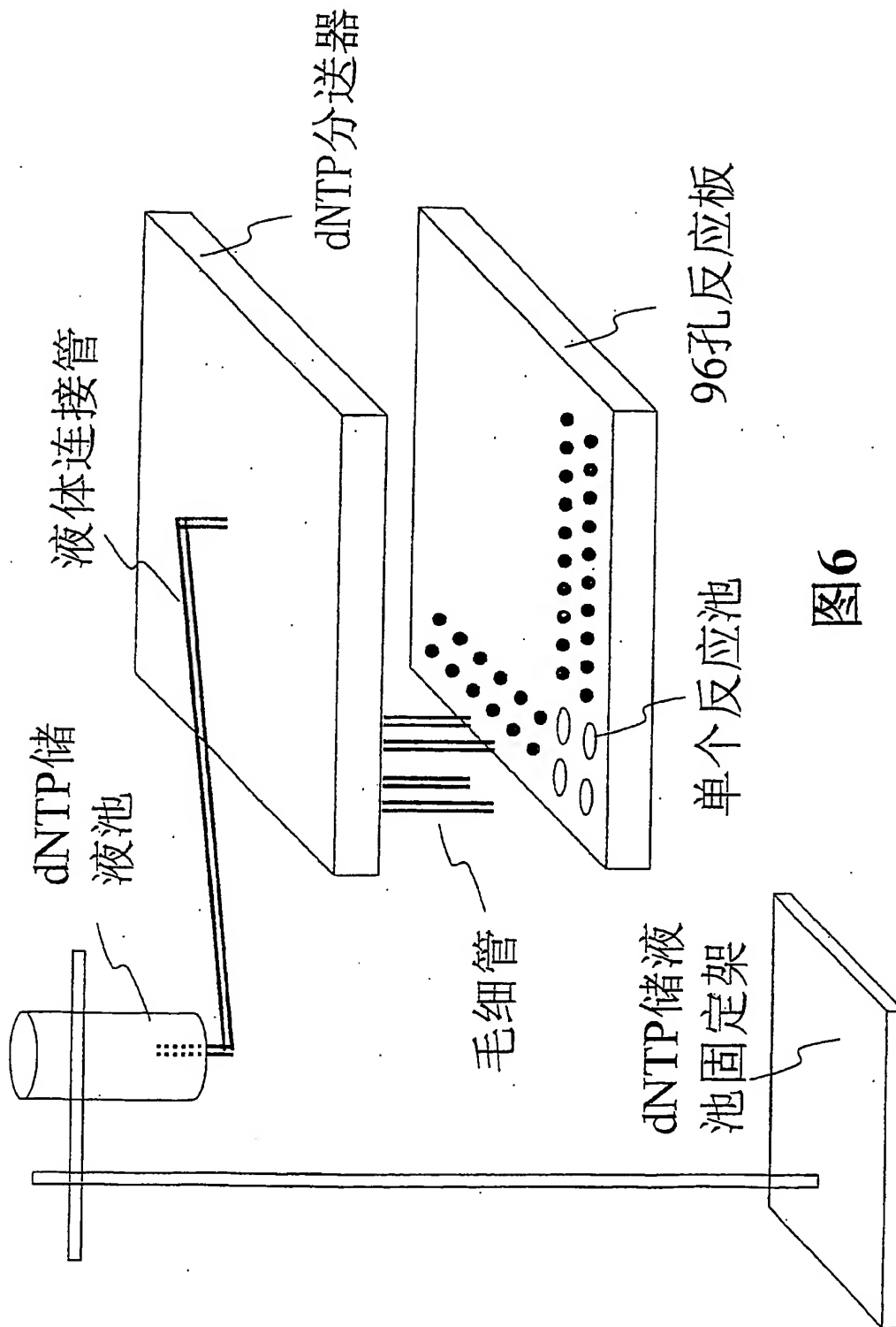


图5



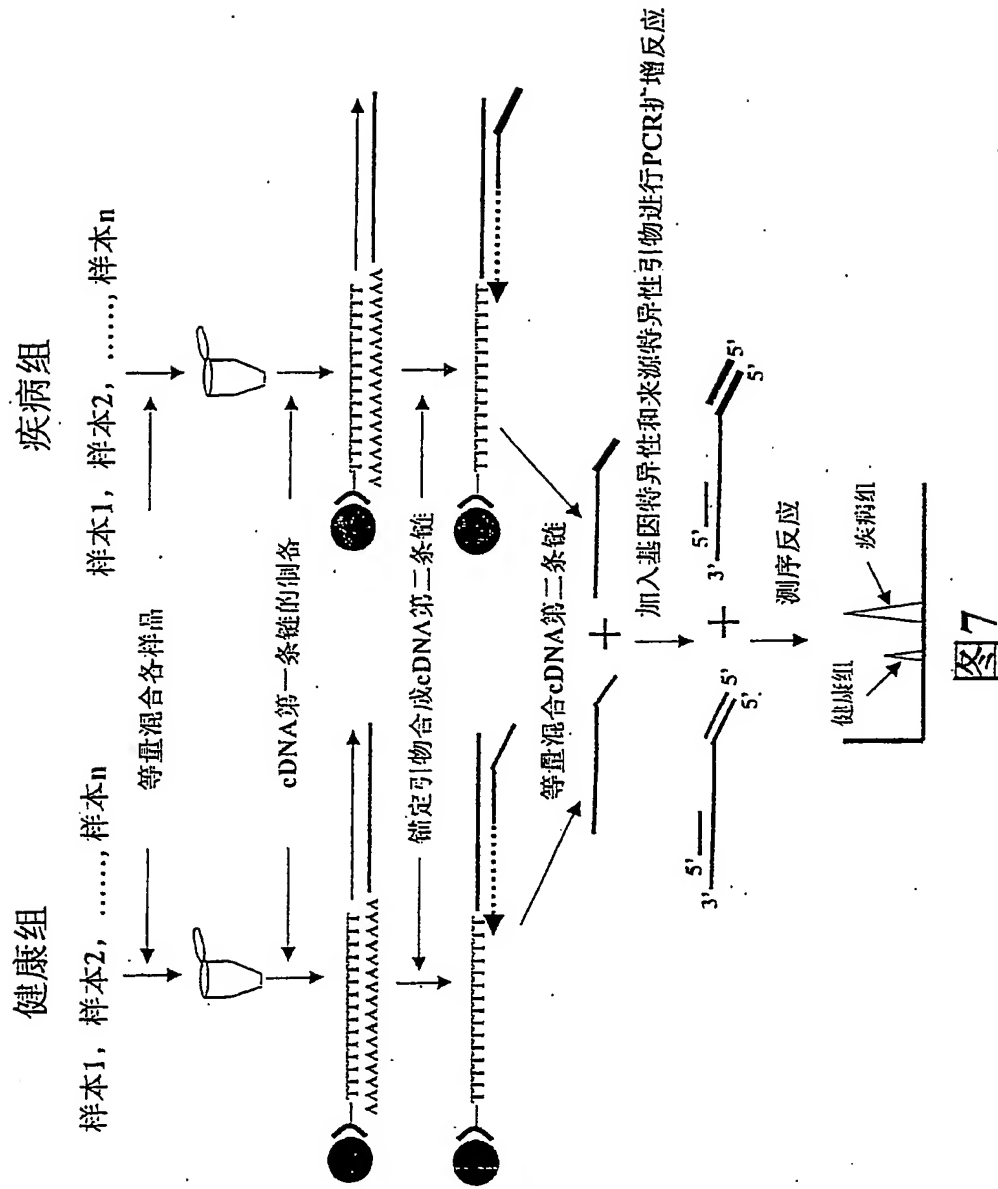


图7

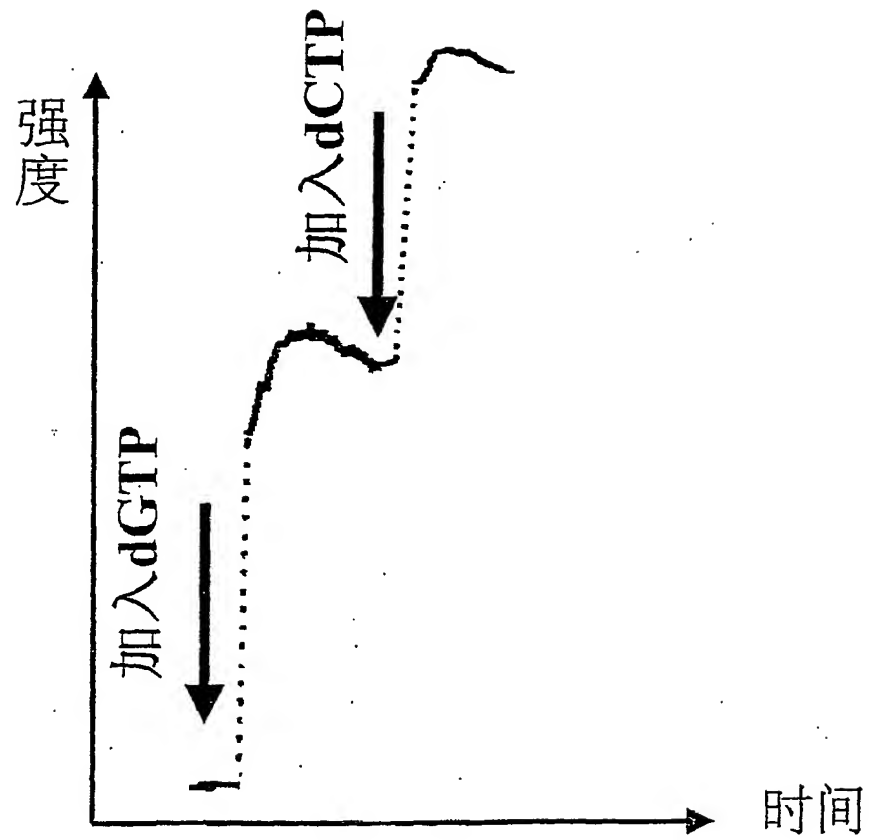


图8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN 03/00648

A. CLASSIFICATION OF SUBJECT MATTER

IPC⁷ C12Q 1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED IPC⁷ C12Q 1/68

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI; EPODOC; PAJ; CNPAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO, A, 9858083 (MEDICAL COLLEGE OHIO) 23.DEC.1998 (1998-12-23) See the whole document	1-10
A	CN, A, 1362527 (UNIV FUDAN) 7.AUG.2002 (2002-08-07) See the whole document	1-10

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
08.November.2003

Date of mailing of the international search report

13 DEC 2003 (18.12.03)

Name and mailing address of the ISA/CN
6 Xitucheng Rd., Jimen Bridge, Haidian District,
100088 Beijing, China
Facsimile No. 86-10-62019451

Authorized officer

Guo wenjie

Telephone No. 86-10-62085346



INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN03/00648

Patent document cited in search report	Publication date	Patent family Members	Publication Date
WO 98/58083 A	23-12-1998	RU 2190019 C2	27-09-2002
		US 5876978 A	02-03-1999
		AU 8072898 A	04-01-1999
		EP 0998581 A2	10-05-2000
		CN 1265155 A	30-08-2002
		JP 2001519672T T	23-10-2001

国际检索报告

国际申请号

PCT/CN 03/00648

A. 主题的分类

IPC⁷ C12 Q1/68

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

IPC⁷ C12 Q1/68

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

WPI; EPODOC; PAJ; CNPAJ

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求编号
A	WO, A, 9858083 (俄亥俄医学院) 1998 年 12 月 23 日 (23.12.98) 参见全文	1-10
A	CN, A, 1362527 (复旦大学) 2002 年 8 月 7 日 (07.08.02)	1-10

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的专用类型:

“A” 明确叙述了被认为不是特别相关的一般现有技术的文件

“E” 在国际申请日的当天或之后公布的在先的申请或专利

“L” 可能引起对优先权要求的怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理

“X” 特别相关的文件, 仅仅考虑该文件, 权利要求所记载的发明就不能认为是新颖的或不能认为是有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 权利要求记载的发明不具有创造性

“&” 同族专利成员的文件

国际检索实际完成的日期

08.11.03

国际检索报告邮寄日期

1 8.12月 2003 (1 8.12.03)

国际检索单位名称和邮寄地址

ISA/CN

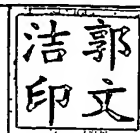
中国北京市海淀区西土城路 6 号(100088)

传真号: 86-10-62019451

受权官员

郭文洁

电话号码: 86-10-62085346



国际检索报告
关于同族专利成员的情报

国际申请号
PCT/CN 03/00648

检索报告中引用的 专利文件	公布日期	同族专利成员	公布日期
WO-A-9858083	23.12 月 1998	RU-C2-2190019	27-09 月 2002
		US-A-5876978	02-03 月 1999
		AU-A- 8072898	04-01 月 1999
		EP-A2-0998581	10-05 月 2000
		CN-A-1265155	30-08 月 2002
		JP-T-2001519672T	23-10 月 2001